

# Hydration of Clays at the Molecular Scale: The Promising Perspective of Classical Density Functional Theory

**Guillaume Jeanmairet<sup>a,b,1</sup>, Virginie Marry<sup>c,d</sup>, Maximilien Levesque<sup>a,b,2</sup>, Benjamin Rotenberg<sup>c,d</sup>, Daniel Borgis<sup>a,b</sup>**

<sup>a</sup>*École Normale Supérieure-PSL Research University, Département de Chimie, 24, rue Lhomond, 75005 Paris, France* <sup>b</sup>: *CNRS, UMR 8640 PASTEUR, F-75005, Paris, France.* <sup>c</sup>: *Sorbonne Universités, UPMC Univ Paris 06, PASTEUR, F-75005, Paris, France.*;

<sup>b</sup>*CNRS UMR 8640 PASTEUR, 75005 Paris, France;*

<sup>c</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR 8234 PHENIX, F-75005, Paris, France;*

<sup>d</sup>*CNRS, UMR 8234 PHENIX, F-75005, Paris, France;*

## Abstract

We report here how the hydration of complex surfaces can be efficiently studied thanks to recent advances in classical molecular density functional theory. This is illustrated on the example of the pyrophyllite clay. After presenting the most recent advances, we show that the strength of this implicit method is that (i) it is in quantitative or semi-quantitative agreement with reference all-atoms simulations (molecular dynamics here) for both the solvation structure and energetics, and that (ii) the computational cost is two to three orders of magnitude less than in explicit methods. The method remains imperfect, in that it locally overestimates the polarization of water close to hydrophilic sites of the clay. The high numerical efficiency of the method is illustrated and exploited to carry a systematic study of the electrostatic and van der Waals components of the surface-solvent interactions within the most popular force field for clays, CLAYFF. Hydration structure and energetics are found to weakly depend upon the electrostatics. We conclude on the consequences of such findings on future force-field development.

---

<sup>1</sup>guillaume.jeanmairet@ens.fr

<sup>2</sup>maximilien.levesque@ens.fr

# 1 Introduction

Pierre Turq, in his recent conference “Histoire d’Eau”, at Université Pierre et Marie Curie, Paris, highlighted the immense challenge that represents modeling of water at the atomic scale in complex, hierarchical materials such as clays[1]. Their surface to volume ratio is high, and have thus key role in various applications such as heterogenous catalysis, electrochemistry, adsorption and transport in porous media. In such materials, being able to model at the same time the smallest and the largest representative scales, *e.g.*, keeping the molecular picture of the solvent while studying mesoscale transport, remains an immense open field. It is our goal to develop tools that will improve the description at the molecular scale, while keeping open the upscaling.

Experimental tools have been developed to get insight into the properties of water at clay surfaces: vibrational sum frequency spectroscopy[2], quasi-elastic neutron scattering techniques[3], infrared (IR) and Raman spectroscopies or nuclear magnetic resonance (NMR). Atomic scale modeling can clarify or give insight into these experimental observations[3, 4, 5, 6]. It was for instance the case of hydrophobic versus hydrophilic behaviours of various clay surfaces and humidities[7, 8] that has been recently clarified by molecular dynamics simulations and theoretical insights[9]. Water dynamics in hectorite clay have also been studied by coupled neutron spin echo and molecular dynamics[3].

This kind of theoretical modeling most often relies on molecular dynamics (MD) and Monte Carlo (MC) simulations, in which each atom is considered individually. Observables are statistical averages over configurations that are obtained deterministically or stochastically. With sufficient sampling, they are theoretically exact. Nevertheless, they are computationally demanding. Nowadays, multiscale materials with several (thousands of) thousands of atoms can be simulated by MD or MC but require using supercomputers and thousands of threads of computational processing units (CPU). At this point, numerical heaviness becomes critical and hinders systematic studies.

To overcome this numerical difficulty, coupled to energetic and economic considerations of running supercomputers, various empirical modeling and physically grounded liquid state theories have emerged. One can for instance forget about the molecular nature of the solvent to only account for a polarizable continuum model (PCM)[10, 11, 12], or simplify the solvent molecule to a dipolar hard sphere particle[13, 14]. Coarse grained strategies have

also been successfully applied to solid-liquid interfaces[15, 16]. Nevertheless, while all these methods proved their efficiency on a panel of problems, the applicability and transferability of their underlying approximations remain to be systematically checked.

These last forty years, liquid state theories[17] have blossomed, classical density functional theory (DFT)[18, 19, 20], integral equations, classical field theories, all have shown to be able to give solvation properties in at least qualitative agreement with reference MD or MC simulations. Nevertheless, they remained toy theories for a long time, *i.e.*, applied only to problems of high symmetry like hard and soft walls. A current challenge lies in the development of a three-dimensional liquid state theory (and its implementation) for complex liquids to describe molecular liquids, solutions, mixtures, in complex environments such as biomolecular media or atomistically resolved surfaces and interfaces. Recent developments in this direction have been done within lattice field theory[21, 22], Gaussian field theory[23], three-dimensional reference interaction site model (3D-RISM)[24, 25, 26, 27, 28, 29]. This last integral equation theory has proved to be useful to a large panel of applications like structure prediction in complex biomolecular systems, but remains difficult to control and improve because of the choice of a closure relation, typically restricted to Hypernetted chain (HNC), Percus-Yevick[30], or Kovalenko-Hirata[28].

Molecular density functional theory (MDFT) has recently been shown to be a powerful tool to study the solvation of complex surfaces and interfaces by a polar solvent, at a fully molecular level of description[31, 32, 33, 34, 35, 36, 37, 38]. In the case of the generic model of dipolar solvent, *i.e.*, the Stockmayer fluid (eventually parameterized to mimic a few properties of water), it was shown to recover quantitatively the structural solvation properties of a complex atomically resolved clay surface, the pyrophyllite, that contains over a thousand atoms[39]. It is the objective of this article to test the most recent developments of MDFT for one of the most important but challenging fluid, namely water, in the typical complex case of the pyrophyllite clay.

In Sec. 2, we present the theoretical and numerical recent developments of MDFT for water. We also briefly introduce our molecular dynamics simulations, that will be used as a reference. In Sec. 3, we apply the method to the solvation of a pyrophyllite clay sheet by water. We investigate both structural and energetic aspects of this hydration. Finally, in Sec. 4, we conclude and give perspectives to the method to overcome actual limitations.

## 2 Theory and Methods

### 2.1 Theory

Molecular density functional theory relies on the definition of a free-energy functional of the six-dimensional position and orientation molecular one-particle density,  $\rho(\mathbf{r}, \mathbf{\Omega})$ , where  $\mathbf{\Omega}$  accounts for the 3 Euler angles. This functional remains unknown and various approximations have been proposed. In the homogeneous reference fluid approximation (HRF), the functional is built from the properties of the bulk solvent. This approach gives accurate results for polar, aprotic fluids[34, 35, 36]. For its part, water requires special attention[38, 37, 40]. We recall here some of the development made to treat water; note that they are also adapted to any solvent with only one Lennard-Jones site. Until now, only simple three-dimensional solutes have been tested in this MDFT framework. Let us define several microscopic densities of a system of  $N$  particles: the position and orientation density of solvent molecules,  $\rho(\mathbf{r}, \mathbf{\Omega})$ , the position density,  $n(\mathbf{r})$ , the charge density,  $\rho_c(\mathbf{r})$ , and the *multipolar* polarization density field,  $\mathbf{P}(\mathbf{r})$ :

$$\rho(\mathbf{r}, \mathbf{\Omega}) = \left\langle \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{\Omega} - \mathbf{\Omega}_i) \right\rangle, \quad (1)$$

$$n(\mathbf{r}) = \left\langle \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle = \int \rho(\mathbf{r}, \mathbf{\Omega}) d\mathbf{\Omega}, \quad (2)$$

$$\rho_c(\mathbf{r}) = \left\langle \sum_{i=1}^N \sigma(\mathbf{r} - \mathbf{r}_i, \mathbf{\Omega}_i) \right\rangle, \quad (3)$$

$$= \iint \sigma(\mathbf{r} - \mathbf{r}', \mathbf{\Omega}) \rho(\mathbf{r}', \mathbf{\Omega}) d\mathbf{r}' d\mathbf{\Omega}, \quad (4)$$

$$\mathbf{P}(\mathbf{r}) = \left\langle \sum_{i=1}^N \boldsymbol{\mu}(\mathbf{r} - \mathbf{r}_i, \mathbf{\Omega}_i) \right\rangle, \quad (5)$$

$$= \iint \boldsymbol{\mu}(\mathbf{r} - \mathbf{r}', \mathbf{\Omega}) \rho(\mathbf{r}', \mathbf{\Omega}) d\mathbf{r}' d\mathbf{\Omega}, \quad (6)$$

where  $\langle \dots \rangle$  indicates an ensemble average and  $d\mathbf{\Omega}$  an angular integration over the three Euler angles, *i.e.*, over all possible orientations of the water molecules and all possible rotations around their  $C_2$  axis. The molecular charge density,  $\sigma(\mathbf{r}, \mathbf{\Omega})$ , and the molecular polarization density field of a

water molecule at the origin,  $\boldsymbol{\mu}(\mathbf{r}, \boldsymbol{\Omega})$ , are:

$$\sigma(\mathbf{r}, \boldsymbol{\Omega}) = \sum_m q_m \delta(\mathbf{r} - \mathbf{s}_m(\boldsymbol{\Omega})), \quad (7)$$

$$\boldsymbol{\mu}(\mathbf{r}, \boldsymbol{\Omega}) = \sum_m q_m \mathbf{s}_m(\boldsymbol{\Omega}) \int_0^1 \delta(\mathbf{r} - u \mathbf{s}_m(\boldsymbol{\Omega})) du, \quad (8)$$

where  $\mathbf{s}_m(\boldsymbol{\Omega})$  is the position of the  $m^{\text{th}}$  (between 1 and 3) atomic site of the water molecule with orientation  $\boldsymbol{\Omega}$ . The molecular charge density and polarization density field are linked by the relation  $\sigma = -\nabla \cdot \boldsymbol{\mu}$ . At first order, the multipolar polarization density field reduces to a molecular dipole located at the origin.

For its part, the solute is described by sites,  $i$ , located at  $\mathbf{R}_i$ , with Lennard-Jones parameters  $\epsilon_i$  and  $\sigma_i$ , and point charges  $q_i$ . The force field associated of the solute is given in Tab. 2. We use the Lorentz-Berthelot mixing rules to determine  $\sigma_{ij}$  and  $\epsilon_{ij}$ . The solute induces a scalar external potential,  $\phi_n$ , and an electric field,  $\mathbf{E}_c$ :

$$\phi_n(\mathbf{r}) = \sum_j 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{|\mathbf{r} - \mathbf{R}_j|} \right)^{12} - \left( \frac{\sigma_{ij}}{|\mathbf{r} - \mathbf{R}_j|} \right)^6 \right], \quad (9)$$

$$\mathbf{E}_c(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_j q_j \frac{\mathbf{r} - \mathbf{R}_j}{|\mathbf{r} - \mathbf{R}_j|^3}, \quad (10)$$

where  $\epsilon_0$  is the vacuum permittivity.

We can extend the original derivation by Evans[18, 19] of the molecular density functional theory to define the solvation free energy functional,  $\mathcal{F}$ , depending of position density and the polarization density.  $\mathcal{F}$  is expressed as the difference in grand potential,  $\Theta - \Theta_b$ , of the system composed of the solute plus the solvent at a given chemical potential, and of the sole homogeneous reference water fluid at bulk density,  $\rho_b = n_b/(8\pi^2)$ , with  $n_b \approx 0.033$  molecules per  $\text{\AA}^3$ :

$$\begin{aligned} \mathcal{F}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] &= \Theta[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] - \Theta_b \\ &= \mathcal{F}_{\text{int}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] + \int \phi(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} - \int \mathbf{E}_c(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}) d\mathbf{r}, \end{aligned} \quad (11)$$

where the intrinsic Helmholtz free energy functional,  $\mathcal{F}_{\text{int}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]$ , can be decomposed into the sum of an ideal part and an excess part,  $\mathcal{F} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{exc}}$ .

Type	O	Si	O	H	Al	H	O	Si	O
$z$ (Å)	39.03	39.62	41.21	41.30	0.00	1.0	1.09	2.68	3.27

Table 1: Coordinates of atomic layers of the pyrophyllite clay along the  $z$  axis, perpendicular to the plane of the layers.

The ideal part is known rigorously:

$$\beta \mathcal{F}_{\text{id}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \iint [\rho(\mathbf{r}, \boldsymbol{\Omega}) \ln \left( \frac{\rho(\mathbf{r}, \boldsymbol{\Omega})}{\rho_b} \right) - \rho(\mathbf{r}, \boldsymbol{\Omega}) + \rho_b] d\mathbf{r} d\boldsymbol{\Omega} \quad (2)$$

where  $\beta \equiv (k_B T)^{-1}$ ,  $k_B$  is the Boltzmann constant, and  $T$  the absolute temperature. On its side, the exact excess term is unknown. Finding a good term for  $\mathcal{F}_{\text{exc}}$  is a current challenge. We proposed recently to expand it around the homogeneous liquid state at the same density,  $\rho_b$ , where we use as sole input the experimental or calculated structure factor,  $S$ , and the experimental longitudinal and transverse electric susceptibilities,  $\chi_L$  and  $\chi_T$ . The reader is referred to Ref. [37] for the complete expression of the excess term of the Helmholtz free energy functional we use therein.

## 2.2 System Description

Pyrophyllite is a neutral clay of space group  $2/m$ , monoclinic, cleaved along orientation  $\{100\}$ . It is a stacking of 9 atomic layers. The top layer is composed of oxygen (O), just below which are silicium (Si) atoms. Both O and Si each form a two-dimensional honeycomb structure. Aluminum (Al) atoms compose the central layer, in 2 thirds of octahedral sites. Between the central layer and the O and Si layers lies an oxygen-hydrogen (O-H) layer. Their, O atoms are at the center of the hexagons formed by Si and top O, with the O-H bond pointing toward the empty octahedral site. Top and side views of pyrophyllite sheets are given in Fig. 1, and coordinates of the atomic layers along the axis perpendicular to the sheet,  $z$ , are given in Tab. 1. We have chosen the central layer as  $z = 0$ .

## 2.3 Numerical Aspects

In both MD and MDFT simulations, atoms of the surface interact with the solvent, water, through optimized Lennard-Jones (van der Waals) and

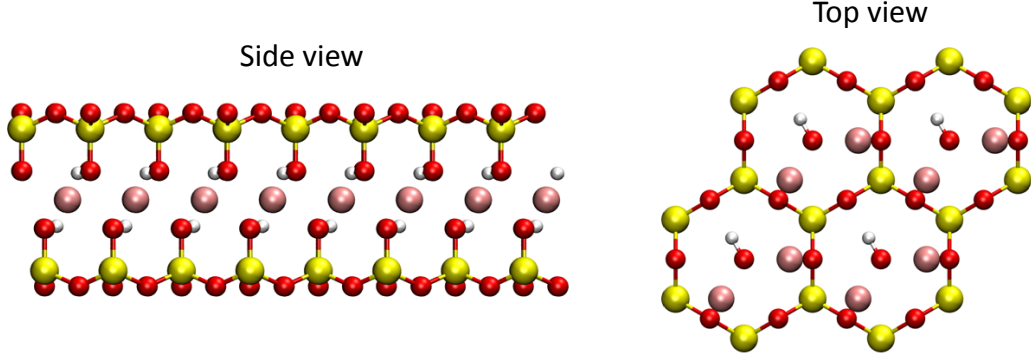


Figure 1: Side and top views of a pyrophyllite sheet. Silicium atoms are in yellow, aluminum in pink, oxygen in red and hydrogen in white.

Molecule	Atom	$\epsilon$ (kJ/mol)	$\sigma$ (Å)	$q$ (e)
Water SPC/E	O	0.65	3.165	-0.8476
	H	0	0	0.4238
Pyrophyllite	Al	5.56388e-6	4.27120	1.575
	Si	7.7005e-6	3.30203	2.1
	O <sub>G</sub>	0.65019	3.16554	-1.050
	O <sub>H</sub>	0.65019	3.16554	-0.950
	H <sub>G</sub>	0.0	0.0	0.425

Table 2: Force field used to model the pyrophyllite and the water solvent. Pyrophyllite parameters are extracted from the CLAYFF force field.

point charges from the force field for clays, CLAYFF[41]. It is a general purpose force field for simulations involving multicomponent minerals, and their interfaces with a solvent. The CLAYFF force field for pyrophyllite is described in Tab. 2.

### 2.3.1 MDFT

The density that minimizes the Helmholtz free energy functional,  $\mathcal{F}$ , defined in Eqs. 11 to 12 is the equilibrium density in the presence of the external potential. In order to find this minimum numerically, we discretize the molecular density,  $\rho(\mathbf{r}, \boldsymbol{\Omega})$ , onto a space and an angular grid. The tridi-

mensional position grid is orthorhombic and has typically 3 to 4 points per Å. The angular grid is of Lebedev type for the two Euler angles describing the spherical orientation of the molecular main axis ( $C_2$  for water). It is a regular, Gauss-Chebyshev quadrature, for the third Euler angle describing the molecular rotation around this main axis. Large numbers of fast Fourier transformed are used to handle convolutions, and this critical part of the algorithm is executed by the Fastest Fourier Transform in the West 3 library (FFTW3). As an optimizer, we use the Limited-Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) quasi-Newton method, as implemented by Byrd, Lu, Nocedal and Zhu[42, 43]. Its main advantage is to only require  $\mathcal{F}$  and its first functional derivative with respect to the density field. The second order derivatives, often needed in Newton-type algorithms, is here approximated with the first derivatives of past iterations. This is a critical advantage for our minimization process.

The external potential,  $\phi_n(\mathbf{r})$ , and the electric field  $\mathbf{E}(\mathbf{r})$  are computed only once, in the earliest steps of the simulation, before the minimization process. We use safe cut-offs of  $4\sigma_i$  for the Lennard-Jones interactions. The electric field is computed by first extrapolating the solute charge density on the grids and then solving the resulting Poisson equation in Fourier space.

In this work, we typically used 64 position nodes per Å<sup>3</sup> and  $\approx 30$  Euler angles per node, for a total of almost 2000 variables to be optimized per Å<sup>3</sup> by L-BFGS. The convergence is reached in approximately 15 iterations, for a convergence criteria of  $10^{-5}$ , as illustrated in Fig. 2. The whole minimization process lasts approximately 25 min on an ordinary laptop, without parallelism. This is *three order of magnitude faster* than reference molecular dynamics.

### 2.3.2 Molecular Dynamics

The simulation box contains 1280 clay atoms, for a dimension  $L_x \times L_y \times L_z = 41.44 \times 35.88 \times 42.30$  Å<sup>3</sup>, which corresponds to 64 elementary, monoclinic, unit cells of formula  $\text{Al}_2[\text{Si}_4\text{O}_{10}](\text{OH})_2$ . The supercell is thus approximately 63 nm<sup>3</sup>. While DFT corresponds to the grand-canonical ensemble, molecular simulations were performed in the canonical ensemble, under conditions such that water far from the surfaces does not feel the effect of the latter. This is achieved in practice by performing prior Monte-Carlo simulations in the NPT ensemble for increasing water contents until the equilibrium distance between the surfaces is sufficient to ensure that the mid-plane density is



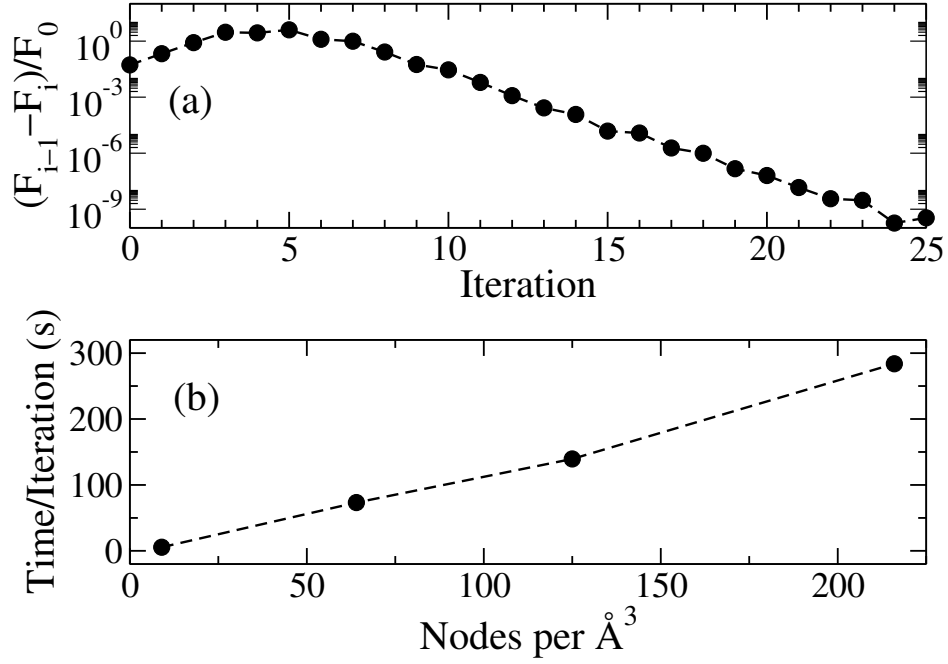


Figure 2: (a) Difference in the solvation free energy functional  $F$  between two iterations normalized by the initial value  $F_0$ , during the iteration process. The grid is composed of  $4 \times 4 \times 4$  nodes per  $\text{\AA}^3$  and 24 discrete orientations of the water molecules per node. A typical minimization is performed with a convergence criteria of  $10^{-5}$  relative variation in free energy, usually reached in 15 iterations. (b) CPU time in seconds per iteration for grid meshes of 3, 4, 5 and 6 nodes per  $\text{\AA}$  and 24 molecular orientations per node. The solvated solute consists in the pyrophyllite sheet described in Section 2.2. The supercell is approximately  $63 \text{ nm}^3$ .

uniform and equal to the density of bulk water. This ensure to set the solvent chemical potential to the value used in MDFT simulation. The obtained configuration has been used as the starting configuration for the molecular dynamics simulation. When simulated by all-atoms molecular dynamics, 1600 water molecules are necessary to fill the pore of pyrophyllite sheets, separated by distance  $L_z$  at 300 K. The temperature is controlled via a Nosé-Hoover thermostat. The simulation package used here is LAMMPS ("Large-scale Atomic/Molecular Massively Parallel Simulator")[44]. After a phase of equilibration, all structural quantities were collected and averaged on a 8 ns trajectory with a 1 fs timestep. Because periodic boundary conditions are applied, the physical system being simulated by MD and MDFT is thus an infinite stacking of infinite sheets.

The local water density,  $n(\mathbf{r})$ , is calculated as the time averaged density in an elementary volume of  $1.0 \text{ \AA}^3$ . The densities in the planes perpendicular to the clay sheet are averaged in the  $z$  planes in an elementary volume of  $0.3 \times L_x \times L_y \text{ \AA}^3$ , where  $L_x$  and  $L_y$  are the length of the supercell in the  $x$  and  $y$  directions. The spatial densities of the average dipoles were calculated in the same way.

### 3 Results

In a previous paper, we have shown that MDFT allows the study of the solvation of complex systems (pyrophyllite, again). This was done with the most simple dipolar solvent: the Stockmayer fluid, a linear molecule modeled by a unique central Lennard-Jones potential associated to two charges separated by a distance that induces the same dipole as water. The functional we used in this previous paper was thus at the dipolar level only. Indeed, as expressed in the previous paper, we did not want to illustrate the difficulties linked to finding the most appropriate functional for water, which is in itself a long term goal, but we wanted to show the efficiency of MDFT in a well-known framework. Here the water molecules can rotate around their  $C_2$  axis. Regarding the functional, it is developed to a fully multipolar level. We first compare the predictions of the molecular density functional theory for the solvent density with our reference all-atom simulations that give exact results. Then, energetics are discussed. We finally analyze the role of electrostatic interactions exerted by the surface on the solvent on these structural and energetic properties.

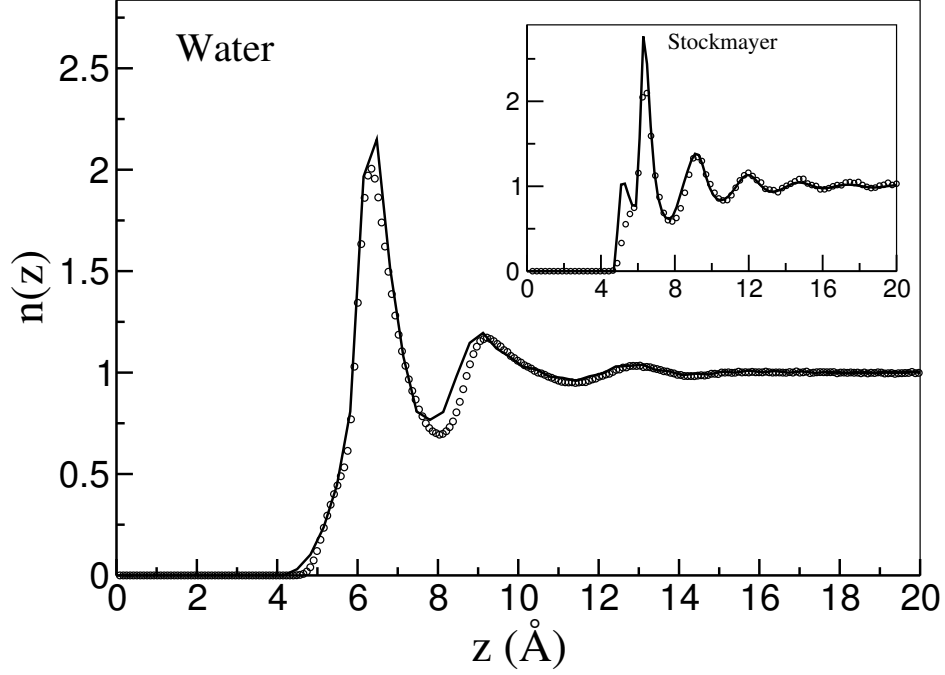


Figure 3: Average density profile of water in the direction  $z$  perpendicular to the pyrophyllite sheet. See Eq. 13 for the definition of  $n_z(z)$ . The profile of reference, calculated by molecular dynamics, is plotted as open circles. MDFT results are shown as a plain lines. The same quantities obtained with the Stockmayer fluid [39] are displayed in the inset.

### 3.1 Density Profiles

The first observable to be computed in a slit pore like the interface of interest here is the averaged number density profile in the direction  $z$  perpendicular to the clay sheet:

$$n_z(z) = \frac{1}{L_x L_y} \iint \frac{n(\mathbf{r})}{n_b} dx dy, \quad (13)$$

where the pyrophyllite sheet is parallel to plane  $(x, y)$ , and the system has size  $L_x \times L_y \times L_z$ . Both MD and MDFT density profiles are given in Fig. 3, where the results obtained for the Stockmayer fluid in Ref. [39] are also plotted.

Importantly, there is a quantitative agreement between the profiles calculated with MDFT and the reference profile.

Three peaks are observed, followed by an homogeneous bulk number den-

sity,  $n(\mathbf{r})/n_b = 1$ . The first peak is broad and high. It is found at  $z = 6.3 \text{ \AA}$ , with  $n_z = 2.2$ . This peak shows a small shoulder at  $z \approx 5.7 \text{ \AA}$ . This shoulder should be reminded later in this work. The first peak corresponds to the first solvation layer of the clay. The second solvation sheet is of relatively small height ( $n_z \approx 1.2$ ). It is found  $2.9 \text{ \AA}$  farther from the clay. A notably depleted zone is found in between the water layers we have just described. An echo of the second water layer is found at  $z = 13.0 \text{ \AA}$ . Its height is almost negligible, 1.03, and would remain unseen without the also almost negligible depletion the precedes it. At higher  $z$  coordinates, the solvent is homogeneous, *i.e.*, is at bulk density  $n_b \approx 0.033$  molecule per  $\text{\AA}^3$ .

All these structural properties are captured by both molecular dynamics and molecular density functional theory, in quantitative agreement, and with the same level of detail. Here, we recover the good agreement between MD and MDFT we found previously for the stockmayer fluid[39], even if the solvent and the functional are much more complicated. It is worth mentioning that when we consider a purely dipolar fluid, we overestimate the prepeak with MDFT. This effect is corrected with better description of the solvent. Nevertheless, by averaging the density  $n(\mathbf{r})$  in the  $(x, y)$  plane, one does not figure out the real *local* water structure on top of the surface.

It is worth emphasizing here that extracting *local* information from MD and MC simulations is time-consuming. Indeed, because of the intrinsic nature of these sampling methods, one cannot compute exactly a local quantity like  $n(\mathbf{r})$ , but rather its convolution by some elementary volume. This means that one effectively measures  $n_z(z) * \Delta z$  or  $n(\mathbf{r}) * \Delta V$ , where  $\Delta z$  and  $\Delta V$  are elementary dimensions chosen arbitrarily and  $*$  denotes the convolution, not  $n_z(z)$  or  $n(\mathbf{r})$ . If the convolution kernels  $\Delta z$  and  $\Delta V$  are too small, the time needed to extract with good statistics the maps shown in Fig. 4, for instance, is increased to a critical point. This problem is *not* relevant in DFT calculations, because it is *not* a method relying on sampling. Indeed, and for the best here, the natural variable in MDFT is the local one-particle density,  $n(\mathbf{r})$ . As an illustration, and in order to get insight to the three dimensional solvation structure of water on top of the clay, we use MDFT. In Fig. 4, we show the tridimensional contour plots of the isosurface of high number densities  $n(\mathbf{r}) = 2$  and 4. The highest densities are found on top of the center of the hexagons, and of silicium atoms of the top-layer (this region is responsible of the prepeak in Fig. 3). Then, a zone of high density forms on top of the hexagone (theses regions are responsible of the main peak in Fig. 3).

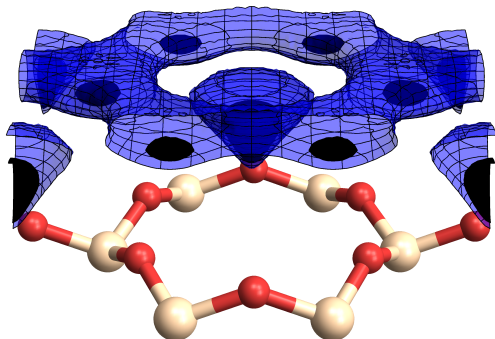


Figure 4: Isosurfaces of density  $n(\mathbf{r}) = 2$  (in blue) and 4 (in black). The honeycomb unit cell is represented for clarity, with oxygen and silicium atoms in red and white, respectively.

Once again, in order to compare MDFT and reference results more locally, we plot in Fig. 5, the normalized density maps in the planes defined by  $z = 5.7, 6.4$  and  $9.2$  Å corresponding to planes of the shoulder, the main and secondary peaks in the density profile of Fig. 3. In the shoulder, the high density zone is strongly localized in the center of the hexagons, really close to the hydrophobic surface. At such close distance from the clay, water is repelled from everywhere but from this tiny zone. Such localization induces that during the averaging leading to Fig. 3, only a shoulder is found. MDFT underestimate the density there, but overestimates the width.

In the first solvation layer, *i.e.*, seen as the main peak of the density profile, water molecules are found on top of the Si atoms. There are  $2.99$  Å in between water molecules. This property is found in both MD and MDFT.

The second solvation layer, *i.e.*, the secondary peak of the density profile, shows high density zones on top of the oxygens and depleted zones on top of Si atoms. There, the variations in number density are not as important as in the previous peaks, but steel between 0.5 and 1.75.

From Fig. 5 it is clear that, despite some disparities, the agreement on density prediction between MD and MDFT is semi-quantitative. Moreover, one can easily see, and we checked it carefully, that convoluting the MDFT map by kernel  $\Delta V$  of length  $1.0$  Å gives back smoothed maps close to the MD maps. The better resolution of MDFT cause theses discrepancies that is why the results are presented in this more meaningful maps, although less comparable. Given its numerical efficiency with respect to MD and its semi-

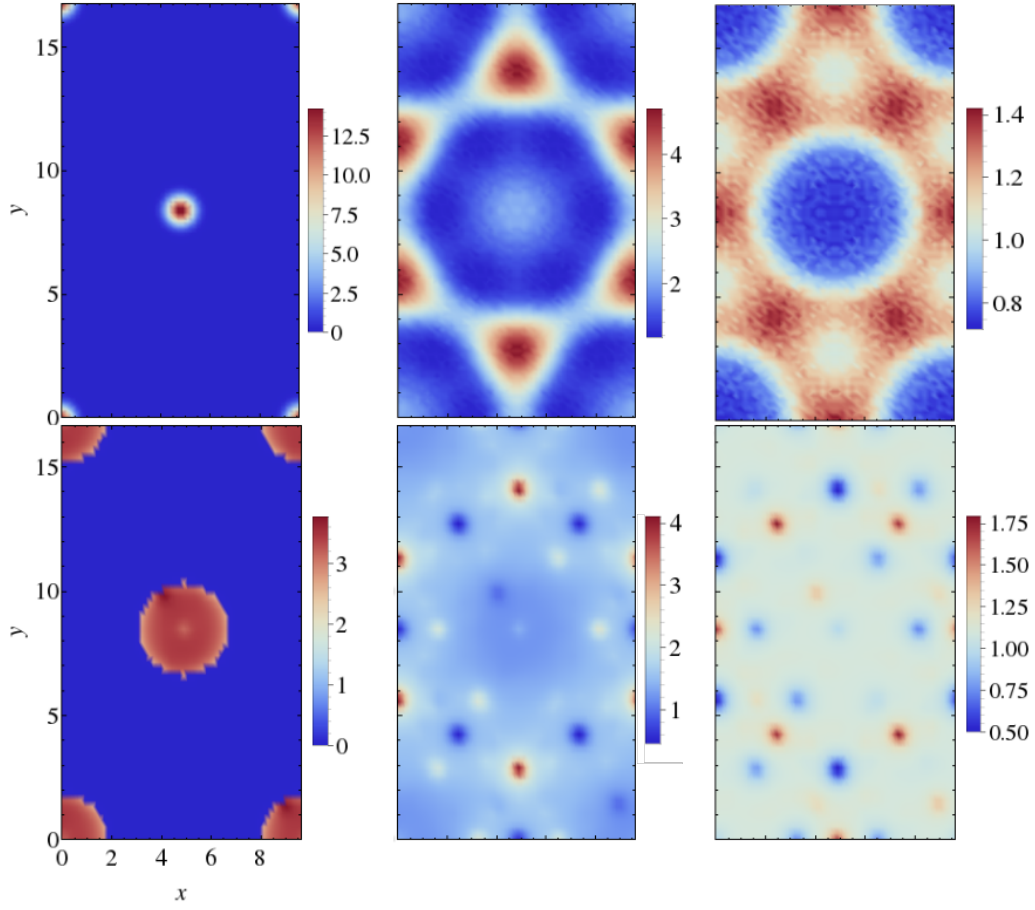


Figure 5: Maps of the normalized number density  $n(\mathbf{r})/n_b$  of water in the planes of the prepeak (left), main peak (center) and secondary peak (right), as calculated by molecular dynamics (top) and molecular density functional theory for water (bottom).

quantitative results, MDFT seems to be a method of choice to address this problem.

### 3.2 Orientational properties

We now turn to the orientational properties of the water molecules that solvate the pyrophyllite surface. In molecular dynamics and Monte Carlo simulations, it is worth emphasizing that orientational properties are difficult to extract, for the same reason as local properties. They are indeed even more local in the phase space. The numerical issue of sufficient sampling of the phase space is recurrent in numerous simulation techniques where orientational degrees of freedom (from electronic spins to molecular orientations) induces new and rich behaviors[45, 46, 47]. For its part, the density functional theory framework and our MDFT implementation give a direct access to this quantity through the full orientational density  $\rho(\mathbf{r}, \mathbf{\Omega})$  defined in Eq. 1. It is even easier to represent orientational observables thanks to the separation of  $\rho$  into the molecular density field,  $n(\mathbf{r})$ , and the polarization density field,  $\mathbf{P}(\mathbf{r})$ , defined in Eq. 2 and 5: They are the natural variables and outputs of MDFT.

In Fig. 6, we report the projections of  $\mathbf{P}$  on the  $z$  axis, noted  $P_z$ , averaged in each plane  $z$ , as a function of  $z$ , as computed by MD and MDFT. The projections on the  $x$  and  $y$  axes have also been computed and found negligible: The maximum values of  $P_x/P_z$  and  $P_y/P_z$  are of the order of  $10^{-3}$ . The polarization density is found to be aligned in the direction perpendicular to the clay sheet, both in MD and MDFT.

Extrema in polarization density correspond to extrema in water density. A first positive peak in  $P_z$  is found at the location of the shoulder in  $n_z$ . Another negative extremum corresponds to the main peak in the density profile. The first polarization extremum is smoother and larger than the second one. Between each maximum, the sign of  $P_z$  changes. The intensity of the macroscopic polarization is linked to the density of molecules and to their orientation in the region of the space considered. Note that since the  $z$  axis is always oriented in the direction leaving the surface, a positive  $P_z$  means that the dipole of the water molecule, pointing from the oxygen atom to the center of mass of the hydrogen atoms, points outward the surface, *i.e.*, that it is the oxygen atom that is the closest to the surface. We can now see that in the first density zone, *i.e.*, in the shoulder, the oxygen atoms of the water molecules are globally the closest to the surface.

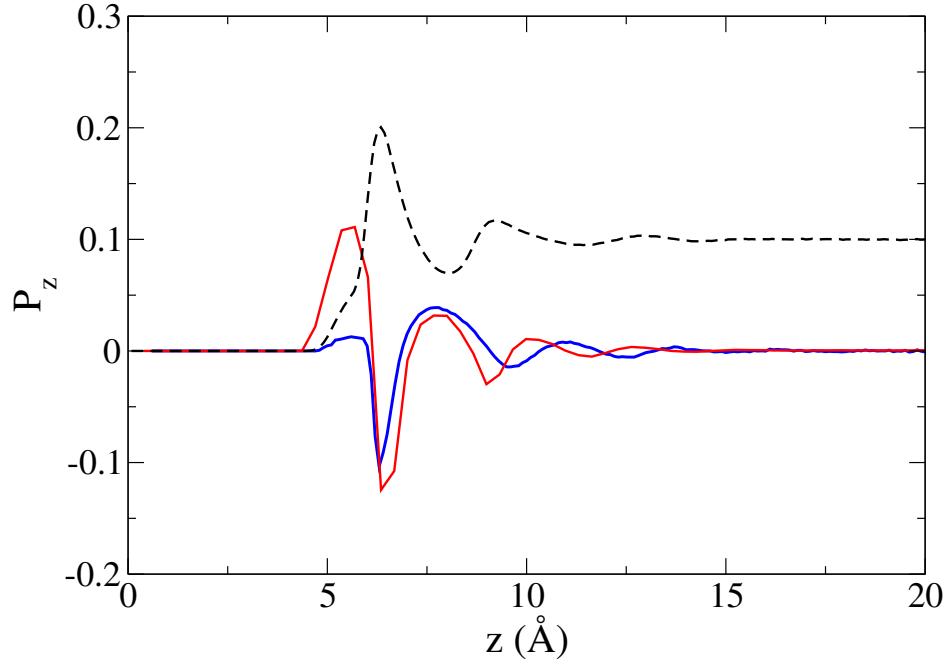


Figure 6: Projection of the polarization on the  $z$  axis,  $P_z$ , in the direction  $z$  perpendicular to the pyrophyllite sheet. The result obtained by MDFT is displayed in red, the one by MD is in blue. The density profile,  $n_z(z)$ , is also shown as a black dashed line, in arbitrary units.



Water molecules on top of Si atoms, (in the main peak), point toward the surface, *i.e.*, hydrogens are globally the closest to the surface. The secondary peak is almost not polarized, there are no preferential orientation of the water molecules in this high density region.

For the Stockmayer fluid, MD and MDFT results were in quantitative agreement. Here, qualitative agreement is found. The polarization profiles obtained by MDFT for water solvent is similar to the one obtained with the Stockmayer fluid[39].

These informations about  $P_z$  highlighted the layer-by-layer structuration of the water molecules, but is not enough to conclude on the importance of orientation. Indeed, a strong polarization can be due either to a small amount of highly oriented molecules or to a large number of water molecules, each having some small preference for a given or several orientations.

To answer this question, we define an orientational order parameters, the local molecular orientation,  $\cos \Theta_\mu(\mathbf{r})$ :

$$\cos \Theta_\mu(\mathbf{r}) = \frac{\|P_z(\mathbf{r})\|}{n(\mathbf{r})}. \quad (14)$$

The average of the local molecular orientation in the plane  $z$ ,  $\cos \Theta_\mu(z) \equiv \langle \cos \Theta_\mu(\mathbf{r}) \rangle_{xy}$ , is plotted in Fig. 7 for MDFT results.

It shows that in the shoulder, that corresponds to this highly dense zone at the center of the hexagons close to the surface, the water molecules are completely oriented. We also showed that this zone have a positive  $P_z$ , which means that water molecules has hydrogen atoms pointing outward the surface. In fact, the oxygen atom of the water molecule is as close as possible of the positively charged silicium atoms, which are just under O atoms of the pyrophyllite. On the opposite the water molecules of the first solvation layer are globally oriented to make the dipoles point toward the surface, but this preferential orientation is very limited as the value of  $\cos \Theta_\mu$  is very weak at the distance corresponding to the first maximum of the density. The second solvation layer is then stacked, translated on top of this first layer and with almost no preferential orientation.

MDFT allow to fully uncover the fine structural solvation properties of pyrophyllite by water.

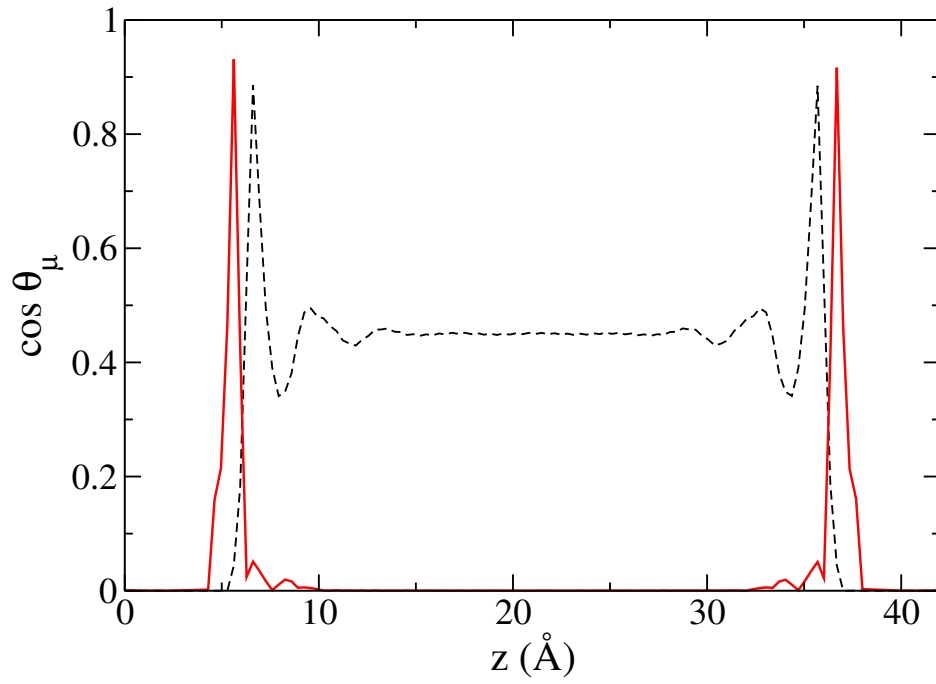


Figure 7: The average molecular orientation,  $\cos \Theta_\mu(z)$ , along the normal to the surface,  $z$ , obtained by MDFT, is shown in full red line. The density profile,  $n_z(z)$ , is shown in black dashed line with arbitrary units to ease the understanding of the position of the peaks.

	$\mathcal{F}$	$\mathcal{F}_{\text{ext}}$	$\mathcal{F}_{\text{id}}$	$\mathcal{F}_{\text{exc}(\text{rad})}$	$\mathcal{F}_{\text{exc}(\text{ori}^\circ)}$
$\gamma_{\text{solv}}$ (mJ/m <sup>2</sup> )	332.4	-12.8	100.4	244.5	0.3

Table 3: The different components of the equilibrium solvation free energy, per surface unit of pyrophyllite. Positive values of  $\gamma_{\text{solv}}$  are meant for hydrophobic surfaces.

### 3.3 Energetics

As shown in Eq. 11, the solvation free energy  $\mathcal{F}$  is decomposed into physically-grounded components: *(i)* the so-called ideal part, that accounts for the configurational entropy, *(ii)* the external part, that accounts for the direct solute-solvent interactions, and *(iii)* the excess part, that accounts for the solvent-solvent correlations, that may be split into a radial and a polarization component. During the minimization process, we can monitor their relative importance. The equilibrium solvation free energy per surface unit of pyrophyllite,  $\gamma_{\text{solv}}$ , is given in Table 3. It is 332.4 mJ/m<sup>2</sup>, a positive value that reflects the known hydrophobic nature of pyrophyllite.

The main contribution to the solvation free energy comes from the radial solvent-solvent interactions, *i.e.*, from the radial part of the excess contribution. The ideal part is also of great importance. Notably, the external and polarization parts do not account much in  $\gamma_{\text{solv}}$ . This must not be understood as a demonstration that polarization and external potentials do not play a critical contribution to the solvation process. Indeed, if one thinks about a hard wall solvated by an ideal gas: the positive infinite external potential inside the wall is the sole cause that induces the zero density inside the wall. Then at equilibrium, the external part of the functional,  $\mathcal{F}_{\text{ext}}$ , is zero, since the density is by nature minimizing its contribution.

### 3.4 The Role of Electrostatics

In a previous paper, we tested the role of the electrostatic interactions between the point charges distributed among atomic sites in the clay within the CLAYFF force field and the solvent molecules. Thanks to the computational efficiency of our MDFT, we were able to systematically study the effect of scaling the CLAYFF point charges from their original values to 0. With the much more realistic picture of water as a solvent, we now do this

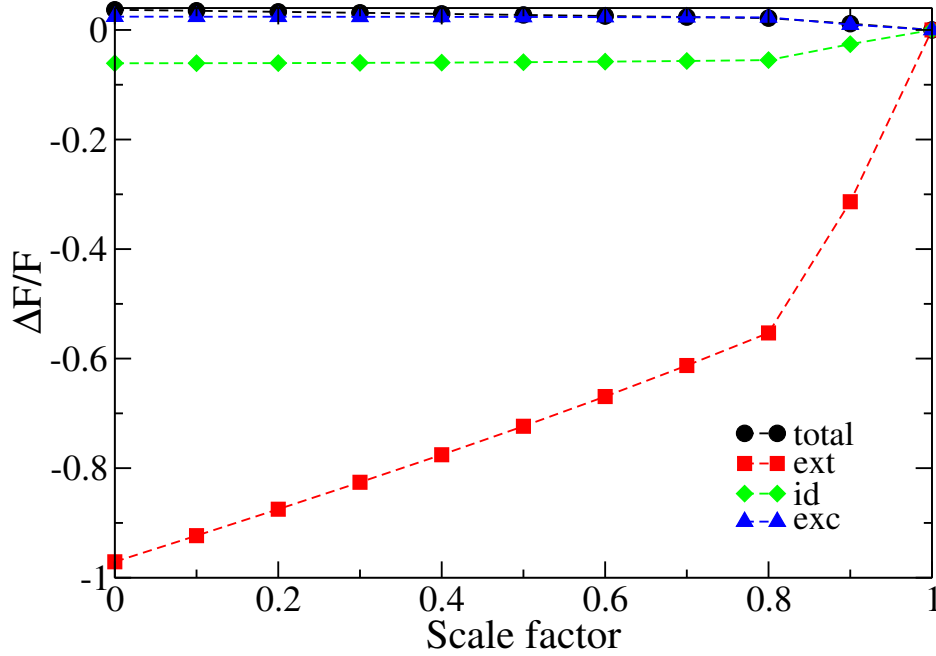


Figure 8: Relative evolution of the various parts of the equilibrium solvation free energy of the clay surface as a function of the solute charge scale factor. For a scale factor of 0, the charges are turned off. It is 1 for the charges directly as proposed by CLAYFF. The relative change in total solvation free energy is 3 %.

same numerical experiment.

The evolution of the hydration free energy with the scale factor is given in Fig. 8. The evolution of the various components of the solvation free energy is also given. The solvation free energy does not change by more than 3 % when the solute charges are turned off. Furthermore, the density profiles are not modified by this process, to a point where plots of  $n_z(z)$  with and without solute charges are indistinguishable.

When the charges are turned off, the total, radial excess and ideal part are almost not changing. The external contribution, on its side, has strong relative changes, but stays small in absolute value. We have confirmation here that the electrostatic contribution of the clay is not important for this system, since turning solute charges off does not change neither the total solvation free energy, nor the solvation structure. Note that for clay minerals

bearing a net total charge (which are in fact more common than electrically neutral ones such as pyrophyllite considered here) with their compensating counter-ions, the effect of polarization could be much more important.

## 4 Conclusion

We have shown that the most recent developments of molecular density functional theory for water allow to study hydration of a complex surface of more than a thousand atoms. As an illustration we studied the pyrophyllite clay. Structural and orientational hydration properties are quantitatively comparable to reference all-atoms simulations, while decreasing the numerical cost by two to three orders of magnitude. The efficiency of MDFT is illustrated by extracting the local structural information in half an hour which otherwise requires several hundred hours of CPU-time with MD. This is due to the intrinsic nature of MDFT, that relies on the local number and polarization densities as natural variables.

We were able to investigate how water solvates the pyrophyllite clay, with (i) a single water molecule on top of the hexagone formed by silicium and oxygen atoms, as close as possible of the plane carrying Si. This single  $\text{H}_2\text{O}$  is strongly oriented with dipole pointing outward from the surface; (ii) then, water molecules form a layer on top of Si atoms and are globally oriented to make their dipoles slightly point toward the surface.

The numerical efficiency of our MDFT approach made it possible to realize a systematic study to analyze the relative contributions of electrostatic and van der Waals interactions, by slowly turning off the charges carried out by the surface. We showed that the density profiles and solvation free energy are not sensitive to the charges of the clay. This may imply that the role of electrostatics in charged clays may be reasonably reduced to their charged defects. This means that the large amount of work dedicated to force field fitting in this field may perhaps be reduced by a focus on important contributions. Such work would rely on a systematic parameters study for which MDFT seems a highly promising tool. This should be done in a future work.

## References

- [1] Meunier, A. *Argiles*; Contemporary publishing international GB science publisher, 2003.
- [2] Richmond, G. L. *Chem. Rev.* **2002**, *102*, 2693–2724.
- [3] Marry, V.; Dubois, E.; Malikova, N.; Durand-Vidal, S.; Longeville, S.; Breu, J. *Env. Sci. & Tech.* **2011**, *45*, 2850–2855.
- [4] Botan, A.; Rotenberg, B.; Marry, V.; Turq, P.; Noetinger, B. *J. Phys. Chem. C* **2011**, *115*, 16109–16115.
- [5] Marry, V.; Rotenberg, B.; Turq, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4802–4813.
- [6] Marry, V.; Dubois, E.; Malikova, N.; Breu, J.; Haussler, W. *J. Phys. Chem. C* **2013**.
- [7] Michot, L. J.; Villieras, F.; Francois, M.; Yvon, J.; Le Dred, R.; Cases, J. M. *Langmuir* **1994**, *10*, 3765–3773.
- [8] Van Oss, C. J.; Giese, R. F. *Clays and Clay Minerals* **1995**, *43*, 474–477.
- [9] Rotenberg, B.; Patel, A. J.; Chandler, D. *J. Am. Chem. Soc.* **2011**, *133*, 20521–20527.
- [10] Leszczynski, J. *Computational Chemistry: Reviews of Current Trends*; World Scientific Publishing Co Pte Ltd, 1999.
- [11] Tomasi, J.; Cammi, R.; Mennucci, B. *Int. J. Quant. Chem.* **1999**, *75*, 783–803.
- [12] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.
- [13] Biben, T.; Hansen, J. P.; Rosenfeld, Y. *Phys. Rev. E* **1998**, *57*, R3727.
- [14] Dzubiella, J.; Hansen, J.-P. *J. Chem. Phys.* **2004**, *121*, 5514.
- [15] Jardat, M.; Dufrêche, J.-F.; Marry, V.; Rotenberg, B.; Turq, P. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2023.
- [16] Jardat, M.; Hribar-Lee, B.; Vlachy, V. *Soft Matter* **2012**, *8*, 954.

- [17] Gray, C. G.; Gubbins, K. E. *Theory of Molecular Fluids: Fundamentals*; Oxford University Press, 1984.
- [18] Evans, R. *Adv. Phys.* **1979**, *28*, 143.
- [19] Evans, R. In Henderson, D., Ed., *Fundamental of Inhomogeneous Fluids*, New York, 1992. Marcel Dekker.
- [20] Wu, J.; Li, Z. *Annu. Rev. Phys. Chem.* **2007**, *58*, 85–112.
- [21] Azuara, C.; Lindahl, E.; Koehl, P.; Orland, H.; Delarue, M. *Nucl. Ac. Res.* **2006**, *34*, W38–W42.
- [22] Azuara, C.; Orland, H.; Bon, M.; Koehl, P.; Delarue, M. *Biophys. J.* **2008**, *95*, 5587–5605.
- [23] Varilly, P.; Patel, A. J.; Chandler, D. *J. Chem. Phys.* **2011**, *134*, 074109.
- [24] Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 7821–7826.
- [25] F. Hirata, E. *Molecular Theory of Solvation*; Kluwer Academic Publishers, Dordrecht, 2003.
- [26] Kloss, T.; Heil, J.; Kast, S. M. *J. Phys. Chem. B* **2008**, *112*, 4337–4343.
- [27] Kloss, T.; Kast, S. M. *J. Chem. Phys.* **2008**, *128*, 134505.
- [28] Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **1998**, *290*, 237–244.
- [29] Yoshida, N.; Imai, T.; Phongphanphanee, S.; Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **2009**, *113*, 873–886.
- [30] Percus, J.; Yevick, G. *Phys. Rev.* **1958**, *110*, 1–13.
- [31] Ramirez, R.; Gebauer, R.; Mareschal, M.; Borgis, D. *Phys. Rev. E* **2002**, *66*.
- [32] Ramirez, R.; Borgis, D. *J. Phys. Chem. B* **2005**, *109*, 6754–6763.
- [33] Ramirez, R.; Mareschal, M.; Borgis, D. *Chem. Phys.* **2005**, *319*, 261–272.
- [34] Gendre, L.; Ramirez, R.; Borgis, D. *Chem. Phys. Lett.* **2009**, *474*, 366–370.

- [35] Zhao, S.; Ramirez, R.; Vuilleumier, R.; Borgis, D. *J. Chem. Phys.* **2011**, *134*, 194102.
- [36] Borgis, D.; Gendre, L.; Ramirez, R. *J. Phys. Chem. B* **2012**, *116*, 2504–2512.
- [37] Jeanmairet, G.; Levesque, M.; Vuilleumier, R.; Borgis, D. *J. Phys. Chem. Lett.* **2013**, *4*, 619–624.
- [38] Jeanmairet, G.; Levesque, M.; Borgis, D. *J. Chem. Phys.* **2013**, *139*, 154101.
- [39] Levesque, M.; Marry, V.; Rotenberg, B.; Jeanmairet, G.; Vuilleumier, R.; Borgis, D. *J. Chem. Phys.* **2012**, *137*, 224107–224107–8.
- [40] Levesque, M.; Vuilleumier, R.; Borgis, D. *J. Chem. Phys.* **2012**, *137*, 034115–034115–9.
- [41] Cygan, R. T.; Liang, J.-J.; Kalinichev, A. G. *J. Phys. Chem. B* **2004**, *108*, 1255–1266.
- [42] Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. *SIAM J. Sci. Comp.* **1995**, *16*, 1190–1208.
- [43] Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. *ACM Trans. Math. Soft.* **1997**, *23*, 550–560.
- [44] Plimpton, S. *Journal of Computational Physics* **1995**, *117*, 1–19.
- [45] Lavrentiev, M. Y.; Nguyen-Manh, D.; Dudarev, S. L. *Phys. Rev. B* **2010**, *81*, 184202.
- [46] Levesque, M.; Martínez, E.; Fu, C.-C.; Nastar, M.; Soisson, F. *Phys. Rev. B* **2011**, *84*, 184205.
- [47] Levesque, M.; Gupta, M.; Gupta, R. P. *Phys. Rev. B* **2012**, *85*, 064111.